IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BOARD OF PATENT APPEALS AND INTERFERENCES

In re patent application of:
Fan, et al. Atty. Docket No.: YOR920030457US1

Serial No.: 10/737,123 Group Art Unit: 2167

Filed: December 16, 2003 Examiner: Le, Miranda

For: SYSTEM AND METHOD FOR ADAPTIVE PRUNING

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPELLANTS' APPEAL BRIEF

Sirs:

Appellant respectfully appeals the final rejection of claims 1-35, in the Office Action dated January 12, 2007. A Notice of Appeal and Pre-Appeal Brief Request was timely filed on April 2, 2007. On February 21, 2008, a Notice of Panel Decision from Pre-Appeal Brief Review was mailed, setting forth a one-month deadline for reply.

I. REAL PARTY IN INTEREST

The real party in interest is International Business Machines Corporation, Armonk, New York, assignee of 100% interest of the above-referenced patent application.

II. RELATED APPEALS AND INTERFERENCES

There are no other appeals or interferences known to Appellants, Appellants' legal representative or Assignee which would directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.

III. STATUS OF CLAIMS

Claims 1-35 are all the claims pending in the application and are under appeal. Claims 1-35 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Venkayala, et al. (U.S. Publication No. 2003/0212679), hereinafter referred to as Venkayala, in view of Rosen, et al. (U.S. Patent No. 6,513,025), hereinafter referred to as Rosen. None of the claims are allowed; all of the rejections are appealed.

IV. STATUS OF AMENDMENTS

In response to the Office Action mailed January 12, 2007 (referred to herein as the "Office Action"), Appellants filed an after-final Response on January 31, 2007. The claims shown in the appendix are shown in their amended form as of the January 31, 2007 Response. On February 21, 2008, a Notice of Panel Decision from Pre-Appeal Brief Review was mailed, setting forth a one-month deadline for reply.

V. SUMMARY OF CLAIMED SUBJECT MATTER

One feature of the invention is a method of searching data in databases using an ensemble of models. Claim 1 defines this feature as follows: "A method of searching data in databases using an ensemble of models." This feature is described at various points in the specification, for example paragraph [0017] describes this feature as

follows: "The use of multiple models (ensembles) can scale up data mining over very large databases and datasets." This is shown in Figure 1.

Another feature of the invention is ordering models within the ensemble in order of prediction accuracy, with the most accurate model being first in the order. Claim 1 defines this feature as follows: "ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order." This feature is described at various points in the specification, for example paragraph [0018] describes this feature as follows: "This training orders models within the ensemble in order of prediction accuracy 110 with the most accurate model being first in the order and joins different numbers of models together to form sub-ensembles 112." This is shown in Figure 1.

Another feature of the invention is selecting a sub-ensemble of the models that meets a given level of confidence, wherein models are joined together in the sub-ensemble in the order of prediction accuracy. Claim 1 defines this feature as follows: "selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Next in the training process, the invention calculates confidence values of each of the sub-ensembles and thereby ranks the sub-ensembles in order of confidence 114." This is shown in Figure 6.

Another feature of the invention is applying the sub-ensemble, in place of the ensemble, to an example to make a prediction. Claim 1 defines this feature as follows: "applying said sub-ensemble, in place of said ensemble, to an example to make a prediction." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "After the training, the invention can make predictions (116, 118, 120) with higher throughput than with the original ensemble." This is shown in Figure 6.

Another feature of the invention is wherein the sub-ensemble includes fewer models than the ensemble. Claim 2 defines this feature as follows: "wherein said sub-

ensemble includes fewer models than said ensemble." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is wherein the confidence is a measure of how closely results from the sub-ensemble will match results from the ensemble. Claim 3 defines this feature as follows: "wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "The confidence is a measure of how closely results from the sub-ensemble will match results from the ensemble." This is shown in Figure 6.

Another feature of the invention is the size of each sub-ensemble is different and has a potentially different level of confidence. Claim 4 defines this feature as follows: "the size of each sub-ensemble is different and has a potentially different level of confidence." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein the size of the ensemble is fixed. Claim 5 defines this feature as follows: "wherein the size of said ensemble is fixed." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in the selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in the selecting process. Claim 6 defines this feature as follows: "wherein as the level of

confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is before the selecting, calculating confidence values of different sub-ensembles. Claim 7 defines this feature as follows: "before said selecting, calculating confidence values of different sub-ensembles." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Next in the training process, the invention calculates confidence values of each of the sub-ensembles and thereby ranks the sub-ensembles in order of confidence 114." This is shown in Figure 6.

Another feature of the invention is a method of searching data in databases using an ensemble of models. Claim 8 defines this feature as follows: "A method of searching data in databases using an ensemble of models." This feature is described at various points in the specification, for example paragraph [0017] describes this feature as follows: "The use of multiple models (ensembles) can scale up data mining over very large databases and datasets." This is shown in Figure 1.

Another feature of the invention is ordering models within the ensemble in order of prediction accuracy, with the most accurate model being first in the order. Claim 8 defines this feature as follows: "ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order." This feature is described at various points in the specification, for example paragraph [0018] describes this feature as follows: "This training orders models within the ensemble in order of prediction accuracy 110 with the most accurate model being first in the order and joins different numbers of models together to form sub-ensembles 112." This is shown in Figure 1.

Another feature of the invention is selecting a sub-ensemble of the models that meets a given level of confidence, wherein models are joined together in the sub-ensemble in the order of prediction accuracy, such that said sub-ensemble include only the most accurate models. Claim 8 defines this feature as follows: "selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy, such that said sub-ensemble include only the most accurate models." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Next in the training process, the invention calculates confidence values of each of the sub-ensembles and thereby ranks the sub-ensembles in order of confidence 114." This is shown in Figure 6.

Another feature of the invention is applying the sub-ensemble, in place of the ensemble, to an example to make a prediction. Claim 8 defines this feature as follows: "applying said sub-ensemble, in place of said ensemble, to an example to make a prediction." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "After the training, the invention can make predictions (116, 118, 120) with higher throughput than with the original ensemble." This is shown in Figure 6.

Another feature of the invention is wherein the sub-ensemble includes fewer models than the ensemble. Claim 9 defines this feature as follows: "wherein said sub-ensemble includes fewer models than said ensemble." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is wherein the confidence is a measure of how closely results from the sub-ensemble will match results from the ensemble. Claim 10 defines this feature as follows: "wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble." This feature is

described at various points in the specification, for example paragraph [0019] describes this feature as follows: "The confidence is a measure of how closely results from the sub-ensemble will match results from the ensemble." This is shown in Figure 6.

Another feature of the invention is the size of each sub-ensemble is different and has a potentially different level of confidence. Claim 11 defines this feature as follows: "the size of each sub-ensemble is different and has a potentially different level of confidence." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein the size of the ensemble is fixed. Claim 12 defines this feature as follows: "wherein the size of said ensemble is fixed." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in the selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in the selecting process. Claim 13 defines this feature as follows: "wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is before the selecting, calculating confidence values of different sub-ensembles. Claim 14 defines this feature as follows: "before said

selecting, calculating confidence values of different sub-ensembles." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Next in the training process, the invention calculates confidence values of each of the sub-ensembles and thereby ranks the sub-ensembles in order of confidence 114." This is shown in Figure 6.

Another feature of the invention is a method of searching data in databases using an ensemble of models. Claim 15 defines this feature as follows: "A method of searching data in databases using an ensemble of models." This feature is described at various points in the specification, for example paragraph [0017] describes this feature as follows: "The use of multiple models (ensembles) can scale up data mining over very large databases and datasets." This is shown in Figure 1.

Another feature of the invention is performing training. Claim 15 defines this feature as follows: "performing training." This feature is described at various points in the specification, for example paragraph [0018] describes this feature as follows: "As show in the flowchart in FIG. 1, the invention first performs training (110, 112, 114)." This is shown in Figure 1.

Another feature of the invention is ordering models within the ensemble in order of prediction accuracy, with the most accurate model being first in the order. Claim 15 defines this feature as follows: "ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order." This feature is described at various points in the specification, for example paragraph [0018] describes this feature as follows: "This training orders models within the ensemble in order of prediction accuracy 110 with the most accurate model being first in the order and joins different numbers of models together to form sub-ensembles 112." This is shown in Figure 1.

Another feature of the invention is joining different numbers of models together to form sub-ensembles, wherein models are joined together in the sub-ensemble in the order of prediction accuracy. Claim 15 defines this feature as follows: "joining different numbers of models together to form sub-ensembles, wherein models are joined together

in said sub-ensemble in said order of prediction accuracy." This feature is described at various points in the specification, for example paragraph [0018] describes this feature as follows: "The models are joined together in the sub-ensemble in the order of prediction accuracy." This is shown in Figure 6.

Another feature of the invention is calculating confidence values of each of the sub-ensembles. Claim 15 defines this feature as follows: "calculating confidence values of each of said sub-ensembles." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Next in the training process, the invention calculates confidence values of each of the sub-ensembles and thereby ranks the sub-ensembles in order of confidence 114." This is shown in Figure 6.

Another feature of the invention is making a prediction. Claim 15 defines this feature as follows: "making a prediction." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "After the training, the invention can make predictions (116, 118, 120) with higher throughput than with the original ensemble." This is shown in Figure 6.

Another feature of the invention is selecting a sub-ensemble of the models that meets a given level of confidence. Claim 15 defines this feature as follows: "selecting a sub-ensemble of said models that meets a given level of confidence." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "First, the invention selects a sub-ensemble that meets a given level of confidence 116." This is shown in Figure 6.

Another feature of the invention is applying the sub-ensemble, in place of the ensemble, to an example to make a prediction. Claim 15 defines this feature as follows: "applying said sub-ensemble, in place of said ensemble, to an example to make a prediction." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "The invention applies the selected sub-ensemble, in place of the ensemble, to an example to make a prediction 118." This is shown in Figure 6.

Another feature of the invention is wherein the sub-ensemble includes fewer models than the ensemble. Claim 16 defines this feature as follows: "wherein said sub-ensemble includes fewer models than said ensemble." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is wherein the confidence is a measure of how closely results form the sub-ensemble will match results from the ensemble. Claim 17 defines this feature as follows: "wherein said confidence is a measure of how closely results form said sub-ensemble will match results from said ensemble." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "The confidence is a measure of how closely results from the sub-ensemble will match results from the ensemble." This is shown in Figure 6.

Another feature of the invention is wherein the size of each sub-ensemble is different and has a potentially different level of confidence. Claim 18 defines this feature as follows: "wherein the size of each sub-ensemble is different and has a potentially different level of confidence." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Thus, the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein the size of the ensemble is fixed. Claim 19 defines this feature as follows: "wherein the size of said ensemble is fixed." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Thus, the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in the selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in the selecting process. Claim 20 defines this feature as follows: "wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is a service of searching data in databases using an ensemble of models. Claim 21 defines this feature as follows: "A service of searching data in databases using an ensemble of models." This feature is described at various points in the specification, for example paragraph [0017] describes this feature as follows: "The use of multiple models (ensembles) can scale up data mining over very large databases and datasets." This is shown in Figure 1.

Another feature of the invention is ordering models within the ensemble in order of prediction accuracy, with the most accurate model being first in the order. Claim 21 defines this feature as follows: "ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order." This feature is described at various points in the specification, for example paragraph [0018] describes this feature as follows: "This training orders models within the ensemble in order of prediction accuracy 110 with the most accurate model being first in the order and joins different numbers of models together to form sub-ensembles 112." This is shown in Figure 1.

Another feature of the invention is selecting a sub-ensemble of the models that meets a given level of confidence, wherein models are joined together in the sub-ensemble in the order of prediction accuracy. Claim 21 defines this feature as follows:

"selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "First, the invention selects a sub-ensemble that meets a given level of confidence 116." This is shown in Figure 6.

Another feature of the invention is applying the sub-ensemble, in place of the ensemble, to an example to make a prediction. Claim 21 defines this feature as follows: "applying said sub-ensemble, in place of said ensemble, to an example to make a prediction." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "The invention applies the selected sub-ensemble, in place of the ensemble, to an example to make a prediction 118." This is shown in Figure 6.

Another feature of the invention is wherein the sub-ensemble includes fewer models than the ensemble. Claim 22 defines this feature as follows: "wherein said sub-ensemble includes fewer models than said ensemble." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is wherein the confidence is a measure of how closely results form the sub-ensemble will match results from the ensemble. Claim 23 defines this feature as follows: "wherein said confidence is a measure of how closely results form said sub-ensemble will match results from said ensemble." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "The confidence is a measure of how closely results from the sub-ensemble will match results from the ensemble." This is shown in Figure 6.

Another feature of the invention is wherein the size of each sub-ensemble is different and has a potentially different level of confidence. Claim 24 defines this feature as follows: "wherein the size of each sub-ensemble is different and has a potentially

different level of confidence." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Thus, the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein the size of the ensemble is fixed. Claim 25 defines this feature as follows: "wherein the size of said ensemble is fixed." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Thus, the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in the selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in the selecting process. Claim 26 defines this feature as follows: "wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is before the selecting, calculating confidence values of different sub-ensembles. Claim 27 defines this feature as follows: "before said selecting, calculating confidence values of different sub-ensembles." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Next in the training process, the invention calculates confidence values of each of the sub-ensembles and thereby ranks the sub-ensembles in order of confidence 114." This is shown in Figure 6

Another feature of the invention is a program storage device readable a computer tangibly embodying a program of instructions executable by the computer for performing a method of searching data in databases using an ensemble of models. Claim 28 defines this feature as follows: "A program storage device readable a computer tangibly embodying a program of instructions executable by said computer for performing a method of searching data in databases using an ensemble of models." This feature is described at various points in the specification, for example paragraph [0017] describes this feature as follows: "The use of multiple models (ensembles) can scale up data mining over very large databases and datasets." This is shown in Figure 1.

Another feature of the invention is ordering models within the ensemble in order of prediction accuracy, with the most accurate model being first in the order. Claim 28 defines this feature as follows: "ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order." This feature is described at various points in the specification, for example paragraph [0018] describes this feature as follows: "This training orders models within the ensemble in order of prediction accuracy 110 with the most accurate model being first in the order and joins different numbers of models together to form sub-ensembles 112." This is shown in Figure 1.

Another feature of the invention is selecting a sub-ensemble of the models that meets a given level of confidence, wherein models are joined together in the sub-ensemble in the order of prediction accuracy. Claim 28 defines this feature as follows: "selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "First, the invention selects a sub-ensemble that meets a given level of confidence 116." This is shown in Figure 6.

Another feature of the invention is applying the sub-ensemble, in place of the ensemble, to an example to make a prediction. Claim 28 defines this feature as follows: "applying said sub-ensemble, in place of said ensemble, to an example to make a

prediction." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "The invention applies the selected sub-ensemble, in place of the ensemble, to an example to make a prediction 118." This is shown in Figure 6.

Another feature of the invention is wherein the sub-ensemble includes fewer models than the ensemble. Claim 29 defines this feature as follows: "wherein said sub-ensemble includes fewer models than said ensemble." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is wherein the confidence is a measure of how closely results form the sub-ensemble will match results from the ensemble. Claim 30 defines this feature as follows: "wherein said confidence is a measure of how closely results form said sub-ensemble will match results from said ensemble." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "The confidence is a measure of how closely results from the sub-ensemble will match results from the ensemble." This is shown in Figure 6.

Another feature of the invention is wherein the size of each sub-ensemble is different and has a potentially different level of confidence. Claim 31 defines this feature as follows: "wherein the size of each sub-ensemble is different and has a potentially different level of confidence." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Thus, the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein the size of the ensemble is fixed. Claim 32 defines this feature as follows: "wherein the size of said ensemble is fixed." This feature is described at various points in the specification, for example paragraph

[0019] describes this feature as follows: "Thus, the size of each of the sub-ensembles is different and has a potentially different level of confidence, while, to the contrary, the size of the ensemble is fixed." This is shown in Figure 6.

Another feature of the invention is wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in the selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in the selecting process. Claim 33 defines this feature as follows: "wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected." This is shown in Figure 6.

Another feature of the invention is before the selecting, calculating confidence values of different sub-ensembles. Claim 34 defines this feature as follows: "before said selecting, calculating confidence values of different sub-ensembles." This feature is described at various points in the specification, for example paragraph [0019] describes this feature as follows: "Next in the training process, the invention calculates confidence values of each of the sub-ensembles and thereby ranks the sub-ensembles in order of confidence 114." This is shown in Figure 6.

Another feature of the invention is a system for searching data in databases using an ensemble of models. Claim 35 defines this feature as follows: "A system for searching data in databases using an ensemble of models." This feature is described at various points in the specification, for example paragraph [0017] describes this feature as follows: "The use of multiple models (ensembles) can scale up data mining over very large databases and datasets." This is shown in Figure 1.

Another feature of the invention is means for ordering models within the ensemble in order of prediction accuracy, with the most accurate model being first in the

order. Claim 35 defines this feature as follows: "means for ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order." This feature is described at various points in the specification, for example paragraph [0018] describes this feature as follows: "This training orders models within the ensemble in order of prediction accuracy 110 with the most accurate model being first in the order and joins different numbers of models together to form sub-ensembles 112." This is shown in Figure 1.

Another feature of the invention is means for selecting a sub-ensemble of the models that meets a given level of confidence, wherein models are joined together in the sub-ensemble in the order of prediction accuracy. Claim 35 defines this feature as follows: "means for selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "First, the invention selects a sub-ensemble that meets a given level of confidence 116." This is shown in Figure 6.

Another feature of the invention is mean for applying the sub-ensemble, in place of the ensemble, to an example to make a prediction. Claim 35 defines this feature as follows: "means for applying said sub-ensemble, in place of said ensemble, to an example to make a prediction." This feature is described at various points in the specification, for example paragraph [0020] describes this feature as follows: "The invention applies the selected sub-ensemble, in place of the ensemble, to an example to make a prediction 118." This is shown in Figure 6.

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

The issues presented for review is whether claims 1-36 are unpatentable under 35 U.S.C. §103(a) by Venkayala, in view of Rosen.

VII. ARGUMENT

A. The Rejection Based on Venkayala and Rosen

1. The Position in the Office Action

The Office Action states:

4. Claims 1-35 are rejected under 35 U.S.C. 103(a) as being unpatentable over Venkayala et a1. (US Pub. No. 20030212679), in view of Rosen et al. (US Pat. No. 6,513,025).

As to claims 1, 28, 35, Venkayala teaches a method of searching data in databases using an ensemble of models (i.e. seeds models, [0018]), said method comprising:

ordering models (i.e. scoring of models, [0017]) within said ensemble in order of prediction accuracy (i.e. Trained model 110 may also be evaluated and adjusted in order to improve the quality, i.e. prediction accuracy, of the model, [0019]), with the most accurate model being first in said order (i.e. a topmost category including a class value having a highest associated probability, [0010]) ([0008-0009, 0018, 0019, 0035]);

selecting a sub-ensemble of said models that meets a given level of confidence (i.e. how much confidence may be placed in the prediction, [0023], the selected class values are those meeting the selection criteria presented in prediction parameters, [0024]), wherein said suberisemble in said order prediction accuracy ([0010, 0035]); and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction (i.e. The selected class values, which are included in multi-category apply output, [0024]) ([0010, 0035]).

Venkayala teaches the step of forming sub-ensembles, wherein said sub ensembles in said order of prediction accuracy (i.e. model apply, [0017]). But Venkayala does not expressly teach "models are joined together in said sub-ensemble".

Rosen teaches "models are joined together in said sub-ensemble" (i.e. The most dependable classification model, that is the classification model associated with the dependability model indicating the highest dependability for the unlabeled example, is chosen to classify each unclassified example (col. 4, line 66 to col. 5, line 9).

It would have been obvious to one of ordinary skill of the art having the teaching of Venkayala and Rosen at the time the invention was made to modify the system of

Venkayala to include wherein models are joined together in said sub-ensemble in said order of prediction accuracy as taught by Rosen.

One of ordinary skill in the art would be motivated to make this combination in order to select the most appropriate classification model, and the prediction of that classification model is then accepted in view of Rosen, as doing so would give the added benefit of coalescing the predictions of the base models by learning the relationships between those predictions and the correct prediction, identifying a specific classification model as the one responsible for producing a final prediction and to including a simple explanation of why the prediction was made, and revising the selected model as the models change over time as taught by Rosen (col. 2, lines 31-42).

As per claim 8, Venkayala teaches a method of searching data in databases using an ensemble of models (i.e. seeds models, [0018]), said method comprising:

ordering models (i.e. scoring of models, [0017]) within said ensemble in order of prediction accuracy (i.e. Trained model 110 may also be evaluated and adjusted in order to improve the quality, i.e. prediction accuracy, of the model, [0019]), with the most accurate model being first in said order (i.e. a topmost category including a class value having a highest associated probability, [0010]) ([0008-0009, 0018, 0019, 0035]);

selecting a sub-ensemble of said models that meets a given level of confidence (i.e. how much confidence may be placed in the prediction, [0023], the selected class values are those meeting the selection criteria presented in prediction parameters, [0024]), wherein said subensemble in said order prediction accuracy, such that said sub-ensemble include only the most accurate models (i.e. a topmost category including a class value having a highest associated probability, [0010]) ([0008, 0009, 0035]); and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction (i.e. The selected class values, which are included in multi-category apply output, [0024]) ([0010, 0035]).

Venkayala teaches the step of forming sub-ensembles, wherein said sub ensembles in said order of prediction accuracy (i.e. model apply, [0017]). Venkayala does not expressly teach "models are joined together in said sub-ensemble".

However, Rosen teaches "models are joined together in said sub-ensemble" (i.e. The most dependable classification model, that is the classification model associated with the dependability model indicating the highest dependability for the unlabeled example, is chosen to classify each unclassified example (col. 4, line 66 to col. 5, line 9).

It would have been obvious to one of ordinary skill of the art having the teaching of Venkayala and Rosen at the time the invention was made to modify the system of Venkayala to include wherein models are joined together in said sub-ensemble in said order of prediction accuracy as taught by Rosen.

One of ordinary skill in the art would be motivated to make this combination in order to select the most appropriate classification model, and the prediction of that classification model is then accepted in view of Rosen, as doing so would give the added benefit of coalescing the predictions of the base models by learning the relationships between those predictions and the correct prediction, identifying a specific classification model as the one responsible for producing a final prediction and to including a simple explanation of why the prediction was made, and revising the selected model as the models change over time as taught by Rosen (col. 2, lines 31-42).

As per claim 15, Venkayala teaches a method of searching data in databases using an ensemble of models (i.e. seeds models, [0018]), said method comprising:

performing training (i.e. training/model building, [0019]) comprising:

ordering models (i.e. scoring of models, [0017]) within said ensemble in order of prediction accuracy (i.e. Trained model 110 may also be evaluated and adjusted in order to improve the quality, i.e. prediction accuracy, of the model, [0019]), with the most accurate model being first in said order (i.e. a topmost category including a class value having a highest associated probability, [0010]) ([0008-0009, 0018, 0019, 0035]);

forming sub-ensembles (i.e. model apply, [0017]), wherein said sub-ensemble in said order of prediction accuracy ([0010, 0035]);

calculating confidence values of each of said sub-ensembles (i.e. to generate one or more scores for each row of data in scoring data. The scores for each row of data

indicate how closely the row of data matches attributes of the model, how much confidence may be placed in the prediction, [0023]); and making a prediction comprising:

selecting a sub-ensemble of said models that meets a given level of confidence (i.e. The selected class values are those meeting the selection criteria presented in prediction parameters, [0024]) ([0010, 0035]); and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction (i.e. The selected class values, which are included in multi-category apply output, [0024]) ([0010, 0035]).

Although Venkayala teaches the step of forming sub-ensembles, wherein said sub ensembles in said order of prediction accuracy (i.e. model apply, [0017]), Venkayala does not expressly teach "joining different number of models together to form sub-ensembles".

Rosen teaches 'joining different number of models together to form sub-ensembles" (i.e. The most dependable classification model, that is the classification model associated with the dependability model indicating the highest dependability for the unlabeled example, is chosen to classify each unclassified example (col. 4, line 66 to col. 5, line 9).

It would have been obvious to one of ordinary skill of the art having the teaching of Venkayala and Rosen at the time the invention was made to modify the system of Venkayala to include joining different numbers of models together to form sub-ensembles, wherein models are joined together in said sub-ensemble in said order of prediction accuracy as taught by Rosen.

One of ordinary skill in the art would be motivated to make this combination in order to select the most appropriate classification model, and the prediction of that classification model is then accepted in view of Rosen, as doing so would give the added benefit of coalescing the predictions of the base models by learning the relationships between those predictions and the correct prediction, identifying a specific classification model as the one responsible for producing a final prediction and to including a simple explanation of why the prediction was made, and revising the selected model as the models change over time as taught by Rosen (col. 2, lines 31-42).

As per claim 21, Venkayala teaches a service of searching data in databases using an ensemble of models (i.e. seeds models, [0018]), said service comprising:

ordering models (i.e. scoring of models, [0017]) within said ensemble in order of prediction accuracy (i.e. Trained model 110 may also be evaluated and adjusted in order to improve the quality, i.e. prediction accuracy, of the model, [0019]), with the most accurate model being first in said order (i.e. a topmost category including a class value having a highest associated probability, [0010]) ([0008-0009, 0018, 0019, 0035]);

selecting a sub-ensemble of said models that meets a given level of confidence (i.e. how much confidence may be placed in the prediction, [0023], the selected class values are those meeting the selection criteria presented in prediction parameters, [0024]), wherein said sub ensemble in said order prediction accuracy ([0010, 0035]); and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction (i.e. The selected class values, which are included in multi-category apply output, [0024]) ([0010, 0035]).

Venkayala teaches the step of forming sub-ensembles, wherein said sub-ensembles in said order of prediction accuracy (i.e. model apply, [0017]). But Venkayala does not expressly teach "models are joined together in said sub-ensemble".

However, Rosen teaches "models are joined together in said sub-ensemble" (i.e. The most dependable classification model, that is the classification model associated with the dependability model indicating the highest dependability for the unlabeled example, is chosen to classify each unclassified example (col. 4, line 66 to col. 5, line 9).

It would have been obvious to one ofordinary skill of the art having the teaching of Venkayala and Rosen at the time the invention was made to modify the system of Venkayala to include wherein models are joined together in said sub-ensemble in said order of prediction accuracy as taught by Rosen.

One of ordinary skill in the art would be motivated to make this combination in order to select the most appropriate classification model, and the prediction of that classification model is then accepted in view of Rosen, as doing so would give the added benefit of coalescing the predictions of the base models by learning the relationships

between those predictions and the correct prediction, identifying a specific classification model as the one responsible for producing a final prediction and to including a simple explanation of why the prediction was made, and revising the selected model as the models change over time as taught by Rosen (col. 2, lines 31-42).

As to claims 2, 9, 16, 22, 29, Venkayala teaches said sub-ensemble includes fewer models than said ensemble (i.e. The selection criterion may comprise one of a topmost category including a class value having a highest associated probability, top N categories including N class values having highest associated probabilities, bottom N categories including N class values having lowest associated probabilities, or a set of select class values specified by the user and their associated probabilities and ranks, [0010]).

As to claims 3, 10, 17, 23, 30, Venkayala teaches said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble (i.e. The scores for each row of data indicate how closely the row of data matches attributes of the model, how much confidence may be placed in the prediction, how likely each output, [0049]).

As to claims 4, 11, 18, 24, 31, Venkayala teaches the size of each sub-ensemble is different and has a potentially different level of confidence (i.e. The selection criterion may comprise one of a topmost category including a class value having a highest associated probability, top N categories including N class values having highest associated probabilities, bottom N categories including N class values having lowest associated probabilities, or a set of select class values specified by the user and their associated probabilities and ranks, [0010]).

As to claims 5, 12, 19, 25, 32, Venkayala teaches the size of said ensemble is fixed (i.e. the selection criteria may include a limit on the number of class values that are to be selected, [0050]).

As to claims 6, 13, 20, 26, 33, Venkayala teaches as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process (i.e. The selection criteria may be defined by desired results data, [0024]), and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in

said selecting process (i.e. The selection criterion may comprise one of a topmost category including a class value having a highest associated probability, top N categories including N class values having highest associated probabilities, bottom N categories including N class values having lowest associated probabilities, or a set of select class values specified by the user and their associated probabilities and ranks, [0010]).

As to claims 7, 14, 27, 34, Venkayala teaches before said selecting, calculating confidence values of different sub-ensembles (i.e. to generate one or more scores for each row of data in scoring data. The scores for each row of data indicate how closely the row of data matches attributes of the model, how much confidence may be placed in the prediction, how likely each output, [0049]).

Response to Arguments

5. Applicant's arguments filed 10/12/06 have been fully considered but they are not persuasive.

A. Applicant argues that the "class values" of Venkayala have nothing to do with a sub ensemble of models. Moreover, the class values are not selected prior to the application of the models, wherein the selected class values are applied in place of the models.

The Examiner respectfully disagrees for the following reasons:

a) The Venkayala teaches "ensemble", and "sub-ensemble", "ordering" as follows:

According to Applicants' Specification, the term "ensemble" and "sub-ensemble" are disclosed as "First the invention performs training. This training *orders models within the ensemble* in order of prediction accuracy with *the most accurate model beingfirst,* and *joined together in the sub-ensemble* in order of prediction accuracy (Instant Specification, Page 1).

Analogously, Venkayala's teaches all such steps as *"generating multi-category* apply output with a plurality of predicted class values and their associated probabilities based on the received input data and a selection criterion" ([0007]), and "The selection criterion may comprise one of *a topmost category including a class value* having a

highest associated probability" top N categories including N class values having highest associated probabilities" ([0008]), wherein:

(a) *multi-category* of Venkayala equates to *the ensemble;*

(b) a *particular category in the multi-category* of Venkayala equates to *the model;*

(c) *a topmost category* of Venkayala equates to *orders models;*

(d) *a top most category* of Venkayala equates to *the most accurate model being first;*

(e) *a topmost category including a class value* of Venkayala equates to *sub-ensemble.*

It is recognized that a class value itself is not a sub-ensemble as Applicants argue; rather, it is *"a topmost category including a class value"* that equivalent to Applicants' *"sub ensemble"*

b) The Venkayala teaches "selection a sub-ensemble" as "The *selection* criterion may comprise one of *a topmost category including a class value"* (See [0007]).

Note that *a topmost category including a class value* of Venkayala equates to *sub ensemble.*

c) The Venkayala teaches "applying said sub-ensemble, *in place of said ensemble.* To an example to make a prediction" as *"A score/prediction* is *a category associated with probability* as the result of applying to *a supervised model* a record whose target value is unknown. A single-target apply operation produces the target value (or category) whose probability is the highest among the all target values" ([0030]), wherein:

a category associated with probability of Venkayala equates to *sub-ensemble. a supervised model a record whose target value is unknown* of Venkayala equates to !!!! *example.*

A score/prediction of Venkayala equates to *make a prediction.*

As mentioned in (a), *a topmost category including a class value (See (e)) is also* go *category associated with probability.*

B. Applicant argues that the "scoring of models" in Venkayala does not teach ordering, ranking, or otherwise sorting models. Instead, the scoring step applies a trained model to make predictions based on data.

The examiner respectfully disagrees.

As aforementioned in (c), *a topmost category ([0008J)* of Venkayala equates to *orders models;*

C. Applicants argue that the generated output having multiple class values has nothing to do with a sub-ensemble of models that is used in place of the ensemble of models. Instead the output 122 having multiple class values is produced after the application of the model 110.

Venkayala teaches *"applying said sub-ensemble, in place of said ensemble, to an example to make a prediction"* as *"A score/prediction* is *a category associated with probability* as the result of applying to *a supervised model* a record whose target value is unknown. A single-target apply operation produces the target value (or category) whose probability is the highest among the all target values" ([0030]) wherein:

a category associated with probability ofVenkayala equates to *sub-ensemble.*

a supervised model a record whose target value is unknown of Venkayala equates to *example.*

A score/prediction ofVenkayala equates to *make a prediction.*

Also *a topmost category including a class value* (See (e)) corresponds to *a category associated with probability*).

D. Applicants argue that the independent claims 1,8, 15,21,28, and 35 defines "selecting a sub-ensemble" (of the ensemble of models) and "applying said sub-ensemble".

Similarly, Venkayala does teach *selecting a sub-ensemble* as "The *selection* criterion may comprise one of *a topmost category including a class value" (See [0007]).*

Note that *a topmost category including a class value* of Venkayala equates to *subensemble.*

Further, Venkayala teach the step of applying to the model after the step of selecting *a topmost category including a class value* as *"A score/prediction* is *a category associated with probability* as the result of applying to *a supervised model* a record whose target value is unknown. A single-target apply operation produces the target value (or category) whose probability is the highest among the all target values" ([0030]), wherein:

a category associated with probability of Venkayala equates to *sub-ensemble.*

a supervised model a record whose target value is unknown of Venkayala equates to !!!! *example.*

A score/prediction of Venkayala equates to *make a prediction.*

And, *a topmost category including a class value (See (e)) is also a category associated with probability.*

Arguments as raised are moot since all claim limitations relevant to this issue have been addressed accordingly.

2. Appellants' Position

a. Independent Claim 1, 8, 15, 21, 28, and 35

Appellants traverse the rejections because the proposed combination of Venkayala and Rosen fails to teach or suggest the claimed features of "applying said sub-ensemble, in place of said ensemble, to an example to make a prediction" as defined in independent claims 1, 8, 15, 21, 28, and 35, nothing within Venkayala discloses ensembles and sub-ensembles of models. Instead, Venkayala teaches that only a <u>single</u> model is built and utilized. More specifically, as illustrated in Fig. 1 of Venkayala, the single trained model 110 is built via training/model building 102, wherein the single trained model 110 is utilized in apply 112 to create scored data 118.. More specifically, the "category associated with probability" of Venkayala (which the Office Action asserts teaches the "sub-ensemble" of the claimed invention) is not applied to the "supervised model" (which the Office Action asserts teaches the "example" of the claimed invention). Rather, the "category associated with probability," in Venkayala, is a result of the

application of "a record whose target value is unknown" to the supervised model. Therefore, as explained in greater detail below, Appellants respectfully submit that the prior art of record does not teach or suggest. The claimed applying of the sub-ensemble to make a prediction.

On page 13, paragraph 1, of the Office Action, the Examiner argues that the "category associated with probability" and "supervised model" of Venkayala teach the "sub-ensemble" and "example", respectively, of the claimed invention. Specifically, the Examiner states that "*category associated with probability* of Venkayala equates to ***sub-ensemble***. *a supervised model a record whose target is unknown* of Venkayala equates to ***an example***" (Office Action, p. 13, para. 1 (emphasis in original)). However, Appellants submit that the "category associated with probability" of Venkayala is not applied to the "supervised model" (independent claims 1, 8, 15, 21, 28, and 35 define "applying said sub-ensemble … to an example to make a prediction"). Instead, in Venkayala, "a record whose target value is unknown" is applied to the "supervised model" in order to obtain the "category associated with probability"(0030 of Venkayala).

In other words, it is the "record whose target value is unknown", and not the "category associated with probability", that is applied to the "supervised model" in Venkayala. In Venkayala, the application of the "record whose target value is unknown" to the "supervised model" results in a "score/prediction", which is a "category associated with probability". Specifically, as described in paragraph 0030 of Venkayala, a score/prediction is a category associated with probability as the result of applying to a supervised model a record whose target value is unknown. A single-target apply operation produces the target value (or category) whose probability is the highest among the all target values. As also discussed in paragraph 0003 of Venkayala, "[o]nce the model is built … [i]t is then used to predict (or score) unknown class values of real-world records". As further discussed in paragraph 0020 of Venkayala, "[i]f the model type is supervised … it can be used to predict (or score) the class value of a record whose class is not known".

To the contrary, as described in paragraph 0020 of Appellants' disclosure, after the training, the invention can make predictions (116, 118, 120) with higher throughput than with the original ensemble. First, the invention selects a sub-ensemble that meets a given level of confidence 116. This level of confidence is supplied by the user through, for example, a graphic user interface or computerized network connection (as discussed below with respect to FIG. 6) and comprises a "means for selecting" a sub-ensemble. As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected. <u>The invention applies the selected sub-ensemble, in place of the ensemble, to an example to make a prediction 118</u>.

Accordingly, contrary to the position taken in the Office Action, the "category associated with probability" of Venkayala is not a sub-ensemble that is applied to the supervised model. Rather, the "category associated with probability" is a result of the application of "a record whose target value is unknown" to the supervised model in Venkayala.

Furthermore, Appellants submit that Rosen is introduced by the Office Action for the limited purpose of illustrating a process of joining models together in a sub-ensemble (Office Action, p. 3, para. 2). Nevertheless, nothing within Rosen mentions applying the sub-ensemble to an example to make a prediction. Therefore, it is Appellants' position that the proposed combination of Venkayala and Rosen fails to teach or suggest the claimed features of "applying said sub-ensemble, in place of said ensemble, to an example to make a prediction" as defined in independent claims 1, 8, 15, 21, 28, and 35.

In addition, the Advisory Action dated February 14, 2007 (referred to herein as the "Advisory Action") argues that the "topmost category" and "received input data" of Venkayala teach the "sub-ensemble" and "example", respectively, of the claimed invention. Specifically, the Advisory Action asserts that "Venkayala discloses 'ensemble' as N categories, and 'sub-ensembles' as 'a topmost category'" (Advisory Action, p. 2, item 1). Moreover, the Advisory Action asserts that "the received input data equates to an example of the claim" (Advisory Action, p. 2, item 2).

However, nothing within Venkayala teaches or suggests that the "topmost category" (which the Advisory Action asserts teaches the "sub-ensemble" of the claimed invention) is applied to the "received input data" (which the Advisory Action asserts teaches the "example" of the claimed invention), in place of the "N categories" (which the Advisory Action asserts teaches the "ensemble" of the claimed invention), to make a prediction. Instead, as described in paragraph 0008 of Venkayala, the "topmost category" of Venkayala merely comprises a portion of a "selection criterion"; and, the "received input data" of Venkayala is used for scoring in a transactional format, wherein input data tables are generated that include active attributes and source attributes. As such, Venkayala fails to teach or suggest the claimed features of "applying said sub-ensemble, in place of said ensemble, to an example to make a prediction" as defined in independent claims 1, 8, 15, 21, 28, and 35.

Furthermore, Appellants draw the Board's attention to paragraph 0021 of Appellants' disclosure, which provides that FIGS. 2 and 3 illustrate the operation of the invention graphically. More specifically, FIG. 2 illustrates the data set for database 200 and the original ensemble 202 of models 204 used to make a prediction 206. FIG. 3 illustrates sub-ensembles 300, 302, 304. Sub-ensemble 300 includes only the most accurate model 308. Sub-ensemble 302 includes the most accurate model 308 and the next most accurate model 310. Sub-ensemble 304 includes the most accurate model 308, the next most accurate model 310, and the third most accurate model 312.

To the contrary, nothing within Venkayala discloses ensembles and sub-ensembles of models. Instead, Venkayala teaches that only a <u>single</u> model is built and utilized. More specifically, as illustrated in Fig. 1 of Venkayala, the single trained model 110 is built via training/model building 102, wherein the single trained model 110 is utilized in apply 112 to create scored data 118.

Nothing within Venkayala teaches or suggests the creation or utilization of ensembles and sub-ensembles of models to make a predication. In Fig. 2 of Venkayala, nothing discloses that a sub-ensemble of models are utilized in item 202; rather, item 202 only discloses the step of "INPUT MODEL", which is singular, not plural. Therefore, it

is Appellants' position that Venkayala fails to teach or suggest the claimed features of "applying said sub-ensemble, in place of said ensemble, to an example to make a prediction" as defined by independent claims 1, 8, 15, 21, 28, and 35. In view the foregoing, the Board is respectfully requested to reconsider and withdraw this rejection.

b. Dependent Claims 2, 9, 16, 22, and 29

Appellants traverse the rejections because the proposed combination of Venkayala and Rosen fails to teach or suggest the claimed features "wherein said sub-ensemble includes fewer models than said ensemble" as defined in dependent claims 2, 9, 16, 22, and 29. As described more fully in section A.2.a, above, nothing within Venkayala discloses ensembles and sub-ensembles of models. Instead, Venkayala teaches that only a single model is built and utilized. Therefore, Appellants submit that Venkayala cannot teach a sub-ensemble of multiple models that includes fewer models than an ensemble of multiple models (dependent claims 2, 9, 16, 22, and 29). In view the foregoing, the Board is respectfully requested to reconsider and withdraw this rejection.

c. Dependent Claims 3, 10, 17, 23, and 30

Appellants traverse the rejections because the proposed combination of Venkayala and Rosen fails to teach or suggest the claimed features "wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble" as defined in dependent claims 3, 10, 17, 23, and 30. As stated on page 9, paragraph 4 – page 10, paragraph 1 of the Office Action, Venkayala discloses that "[t]he scores for each row of data indicate how closely the *row of data* matches attributes of the *model*" (citing Venkayala, para. 0049 (emphasis added)). Thus, the Office Action asserts that the "row of data" and "model" of Venkayala teach the "sub-ensemble" and "ensemble of models" of the claimed invention.

Nevertheless, nothing within Venkayala teaches or suggests that the "row of data" of Venkayala (which the Office Action asserts teaches the "sub-ensemble") is used in place of the "model" to make a prediction (independent claims 1, 8, 15, 21, 28, and 35

define "applying said sub-ensemble, in place of said ensemble, to an example to make a prediction"). Instead, both the "row of data" and the "model" of Venkayala are utilized to make predictions.

More specifically, as illustrated in FIG. 1 of Venkayala, both the scoring data 116 (which comprises the "row of data") and the trained model 110 are utilized in the apply step 112. As described in paragraph 0020 of Venkayala, the apply step 112 involves using the deployed trained model 110 to make predictions or recommendations based on new input data for scoring 116.

Accordingly, Appellants submit that the "row of data" of Venkayala (which the Office Action asserts teaches the "sub-ensemble") is not used in place of the "model" to make a prediction. Instead, both the "row of data" and the "model" of Venkayala are utilized to make predictions. Furthermore, Appellants submit that Rosen is introduced by the Office Action for the limited purpose of illustrating a process of joining models together in a sub-ensemble (Office Action, p. 3, para. 2). Nevertheless, nothing within Rosen mentions measuring how closely results of the sub-ensemble matches results from an ensemble of models. Therefore, it is Appellants' position that the proposed combination of Venkayala and Rosen fails to teach or suggest the claimed features "wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble" as defined in dependent claims 3, 10, 17, 23, and 30.

d. Dependent Claims 6, 13, 20, 26, and 33

Appellants traverse the rejections because the proposed combination of Venkayala and Rosen fails to teach or suggest the claimed features "wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process" as defined in dependent claims 6, 13, 20, 26, and 33.

As discussed above in section A.2.c, the Office Action argues that Venkayala discloses the claimed features "wherein said *confidence* is a measure of how closely

results from said *sub-ensemble* will match results from said ensemble" (dependent claims 3, 10, 17, 23, and 30 (emphasis added)). Specifically, the Office Action asserts that Venkayala discloses that "[t]he *scores* for each row of data indicate how closely the *row of data* matches attributes of the model" (Office Action, p. 9, para. 4 – p. 10, para. (citing Venkayala, para. 0049 (emphasis added))). Thus, the Office Action asserts that the "scores" and "row of data" of Venkayala teach the "confidence" and "sub-ensemble", respectively, of the claimed invention. Nevertheless, Appellants submit that the "row of data" of Venkayala does not include any models; and as such, Venkayala does not select a "row of data" based on the number of models in the row. Instead, as illustrated in Figure 1 of Venkayala, the "row of data" (which is included in the scoring data 116) is utilized with the model 110 in an apply operation 112 to output scored data 118.

In other words, Venkayala does not select a "row of data" (which the Office Action asserts teaches the "sub-ensemble" of the claimed invention) based on the number of models included in the "row of data" because: (a) Venkayala does not select a "row of data" to apply to an example in place of an ensemble; and, (b) the "row of data" does not include any models. Rather, the "row of data" is utilized with the model 110 in an apply operation 112 in order to produce scored data 116.

Furthermore, in its rejection of dependent claims 6, 13, 20, 26, and 33, the Office Action references Venkayala's discussion of highest and lowest "associated probabilities" (Office Action, p. 10, para. 4 – p. 11, para. 1). However, the "associated probabilities" of Venkayala are unrelated to the "level of confidence" of the claimed invention (dependent claims 6, 13, 20, 26, and 33 define "wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process"). More specifically, the "level of confidence" of the claimed invention "is a measure of how closely results from said sub-ensemble will match results from said ensemble" (dependent claims 3, 10, 17, 23, and 30).

Conversely, the “associated probabilities” of Venkayala measure of how accurately the apply step 112 predicted the class values. In other words, Venkayala outputs predicted class values; and, the “associated probabilities” of Venkayala measure the prediction accuracy of the class values (Venkayala, para. 0007). Nevertheless, the “associated probabilities” of Venkayala are unrelated to matching results of a sub-ensemble and an ensemble. Therefore, the “associated probabilities” of Venkayala are unrelated to the “level of confidence” of the claimed invention.

As described in paragraph 0020 of the claimed invention, after the training, the invention can make predictions (116, 118, 120) with higher throughput than with the original ensemble. First, the invention selects a sub-ensemble that meets a given level of confidence 116. This level of confidence is supplied by the user through, for example, a graphic user interface or computerized network connection (as discussed below with respect to FIG. 6) and comprises a "means for selecting" a sub-ensemble. As the level of confidence is raised, a sub-ensemble that has more models will be selected and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected. The invention applies the selected sub-ensemble, in place of the ensemble, to an example to make a prediction 118.

Accordingly, Appellants submit that the “row of data” of Venkayala (which the Office Action asserts teaches the “sub-ensemble”) does not include any models; and as such, Venkayala does not select a “row of data” based on the number of models in the row. Instead, as illustrated in Figure 1 of Venkayala, the “row of data” (which is included in the scoring data 116) is utilized with the model 110 in an apply operation 112 to output scored data 118. Furthermore, Appellants submit that Rosen is introduced by the Office Action for the limited purpose of illustrating a process of joining models together in a sub-ensemble (Office Action, p. 3, para. 2). Nevertheless, nothing within Rosen mentions selecting sub-ensembles based on a confidence level of the sub-ensembles and the number of models included in the sub-ensembles. Therefore, it is Appellants’ position that the proposed combination of Venkayala and Rosen fails to teach or suggest the claimed features “wherein as the level of confidence is raised, a sub-

ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process" as defined in dependent claims 6, 13, 20, 26, and 33.

e. Dependent Claims 4, 5, 7, 11, 12, 14, 18, 19, 24, 25, 27, 31, 32, and 34

It is Appellants' position that the proposed combination of Venkayala and Rosen does not render obvious independent claims 1, 8, 15, 21, and 28 and similarly does not render obvious dependent claims 4, 5, 7, 11, 12, 14, 18, 19, 24, 25, 27, 31, 32, and 34. In view the foregoing, the Board is respectfully requested to reconsider and withdraw this rejection.

B. CONCLUSION

In view the forgoing, the Board is respectfully requested to reconsider and withdraw the rejections of claims 1-35.

Please charge any deficiencies and credit any overpayments to Attorney's Deposit Account Number 50-0510.

Respectfully submitted,

Date: March 19, 2008

/Duane N. Moore/
Duane N. Moore, Esq.
Registration No. 53,352

Gibb & Rahman, LLC
2568-A Riva Road, Suite 304
Annapolis, MD, 21401
Voice: (410) 573-6501
Fax: (301) 261-8825
Customer No. 29154

IX. CLAIMS APPENDIX

1. A method of searching data in databases using an ensemble of models, said method comprising:

ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy; and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

2. The method in claim 1, wherein said sub-ensemble includes fewer models than said ensemble.

3. The method in claim 1, wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble.

4. The method in claim 1, wherein, the size of each sub-ensemble is different and has a potentially different level of confidence.

5. The method in claim 1, wherein the size of said ensemble is fixed.

6. The method in claim 1, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.

7. The method in claim 1, further comprising, before said selecting, calculating confidence values of different sub-ensembles.

8. A method of searching data in databases using an ensemble of models, said method comprising:

ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy, such that said sub-ensemble include only the most accurate models; and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

9. The method in claim 8, wherein said sub-ensemble includes fewer models than said ensemble.

10. The method in claim 8, wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble.

11. The method in claim 8, wherein the size of each sub-ensemble is different and has a potentially different level of confidence.

12. The method in claim 8, wherein the size of said ensemble is fixed.

13. The method in claim 8, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.

14. The method in claim 8, further comprising, before said selecting, calculating confidence values of different sub-ensembles.

15. A method of searching data in databases using an ensemble of models, said method comprising:

performing training comprising:

ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

joining different numbers of models together to form sub-ensembles, wherein models are joined together in said sub-ensemble in said order of prediction accuracy;

calculating confidence values of each of said sub-ensembles; and

making a prediction comprising:

selecting a sub-ensemble of said models that meets a given level of confidence; and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

16. The method in claim 15, wherein said sub-ensemble includes fewer models than said ensemble.

17. The method in claim 15, wherein said confidence is a measure of how closely results form said sub-ensemble will match results from said ensemble.

18. The method in claim 15, wherein the size of each sub-ensemble is different and has a potentially different level of confidence.

19. The method in claim 15, wherein the size of said ensemble is fixed.

20. The method in claim 15, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.

21. A service of searching data in databases using an ensemble of models, said service comprising:

ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy; and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

22. The service in claim 21, wherein said sub-ensemble includes fewer models than said ensemble.

23. The service in claim 21, wherein said confidence is a measure of how closely results form said sub-ensemble will match results from said ensemble.

24. The service in claim 21, wherein the size of each sub-ensemble is different and has a potentially different level of confidence.

25. The service in claim 21, wherein the size of said ensemble is fixed.

26. The service in claim 21, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level

of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.

27. The service in claim 21, further comprising, before said selecting, calculating confidence values of different sub-ensembles.

28. A program storage device readable a computer tangibly embodying a program of instructions executable by said computer for performing a method of searching data in databases using an ensemble of models, said method comprising:

ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy; and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

29. The program storage device in claim 28, wherein said sub-ensemble includes fewer models than said ensemble.

30. The program storage device in claim 28, wherein said confidence is a measure of how closely results form said sub-ensemble will match results from said ensemble.

31. The program storage device in claim 28, wherein the size of each sub-ensemble is different and has a potentially different level of confidence.

32. The program storage device in claim 28, wherein the size of said ensemble is fixed.

33. The program storage device in claim 28, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.

34. The program storage device in claim 28, further comprising, before said selecting, calculating confidence values of different sub-ensembles.

35. A system for searching data in databases using an ensemble of models, said method comprising:

means for ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

means for selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy; and

means for applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

X. EVIDENCE APPENDIX

There is no other evidence known to Appellants, Appellants' legal representative or Assignee which would directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.

XI. RELATED PROCEEDINGS APPENDIX

There is no other related proceedings known to Appellants, Appellants' legal representative or Assignee which would directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.